WHAT IS CLAIMED IS:

1.   A form sheet type determining method comprising the steps of:

extracting each character string on an input form sheet as a keyword, after performing character recognition on the each character string; and

collating the extracted keywords with a plurality of sets of keywords registered beforehand for each predetermined form sheet as one set of keywords in a keyword register, thereby to determine the type of said input form sheet.

2.   A method according to claim 1, wherein each keyword in each set of keywords registered beforehand is registered in said keyword register in association with a predetermined corresponding weight, and

wherein in said step of collating, each of said extracted keywords of said input form sheet is given a weight; the degree of matching between said input form sheet and said predetermined form sheet types is evaluated for each predetermined form sheet type by using said weights of said extracted keywords and said predetermined weights of the keywords in each set of said form sheet types within said keyword register; and one of said predetermined form sheet types having the highest degree of matching is determined to be the type of the input form sheet.

3.   A method according to claim 2, wherein said

predetermined weight of each keyword of said sets of keywords registered beforehand is a keyword-specific weight.

4.   A method according to claim 2, wherein the weights attached to each of said extracted keywords of said input form sheet includes at least a weight based on the type of characters forming the keyword and a weight based on the location of the keyword on said input form sheet.

5.   A form sheet type determining method for determining to which of predetermined form sheet types an input form sheet corresponds, comprising the steps of:

        registering a plurality of sets of keywords beforehand in a keyword register with one set of keywords for each of predetermined form sheet types;

        reading image data of an input form sheet, extracting character strings from the read image data, and performing character recognition on each of the extracted character strings;

        extracting each of said character-recognized character strings as a keyword;

        collating said extracted keywords, for each of the form sheet types, with said plurality of sets of keywords registered in said register, there by to determine the type of said input form sheet.

6.   A method according to claim 5, wherein in said keyword register, said each keyword in said sets

of keywords is registered in association with a predetermined corresponding weight, and

wherein in said step of collating, each of said extracted keywords of said input form sheet is attached with a weight; the degree of matching between said input form sheet and said predetermined form sheet types is evaluated for each predetermined form sheet type by using said weights of said extracted keywords and said predetermined weights of the keywords in each set of said form sheet types within said keyword register; and one of said predetermined form sheet types having the highest degree of matching is determined to be the type of the input form sheet.

7. A method according to claim 6, wherein the weight attached to each of said extracted keywords of said input form sheet is a weight based on the type of characters forming the keyword.

8. A method according to claim 6, wherein the weight attached to each of said extracted keywords of said input form sheet is a weight based on the location on said input form sheet.

9. A method according to claim 6, wherein said predetermined weight of each keyword of said registered set of keywords is a keyword-specific weight.

10. A method according to claim 8, wherein the weight attached to each of said extracted keywords of said input form sheet based on the location on said form sheet, is given a larger weight as the location of

the keyword on the input form sheet approaches closer
to the uppermost location.

11.   A method according to claim 6, wherein the
weights attached to each of said extracted keywords of
said input form sheet include a weight based on the
type of characters forming the keyword and a weight
based on the location of the keyword on said input form
sheet.

12.   A method according to claim 5 further
comprising a step of forming one or more new keywords
by taking out arbitrary two or more keywords from the
extracted keywords extracted in said extracting step,
and by combining the taken out keywords, and

        in said step of collating, said extracted
keywords and said formed new one or more keywords are
collated, for each of the form sheet types, with said
sets of keywords registered in said keyword register,
thereby to determine the type of said input form sheet.

13.   A form sheet type determining method for
determining to which of predetermined form sheet types
an input form sheet corresponds, comprising the steps
of:

        registering a plurality of sets of keywords
beforehand in a keyword register with one set of
keywords for each of predetermined form sheet types;

        reading image data of an input form sheet,
extracting character strings from the read image data,
and performing character recognition on each of the

extracted character strings;

collating said character-recognized character strings with reference character string patterns stored in a data base beforehand, and extracting as a keyword each of the character strings which coincide at least partly with an arbitrary one of the reference character patterns;

collating said extracted keywords, for each of the form sheet types, with said sets of keywords registered in said register, thereby to determine the type of said input form sheet.

14.   A method according to claim 13 further comprising a step of forming one or more new keywords by taking out arbitrary two or more keywords from the extracted keywords extracted in said extracting step, and by combining the taken out keywords, and

in said step of collating, said extracted keywords and said formed new one or more keywords are collated, for each of the form sheet types, with said sets of keywords registered in said keyword register, thereby to determine the type of said input form sheet.

15.   A form sheet type determining apparatus for determining to which of predetermined form sheet types an input form sheet corresponds, comprising:

a keyword register which stores therein a plurality of sets of keywords one set for each of predetermined form sheet types;

a character recognition unit which reads

image data of an input form sheet, extracts character

strings from the read image data, and performs

character recognition on each character string

extracted;

a keyword extraction unit which extracts as a

keyword each of the character strings character-

recognized by the character recognition unit;

a collator which collates said extracted

keywords, for each predetermined form sheet type, with

each set of keywords of said plurality of sets of

keywords registered in said keyword register to thereby

determine the type of said input form sheet.

16.   An apparatus according to claim 15, wherein

in said collator each of said extracted keywords is

given a weight based on a type of characters

constituting the extracted further keyword.

17.   An apparatus according to claim 16, wherein

said type of characters distinguishes whether each of

said extracted keywords is typed one or handwritten one.

18.   An apparatus according to claim 15, wherein

in said collator each of said extracted keywords is

given a weight in accordance with a location of the

keyword on said input form sheet.

19.   An apparatus according to claim 15, wherein

in said register each keyword in each set of keywords

is registered in association with a corresponding

keyword-specific weight for each of form sheet types.

20.   An apparatus according to claim 15, wherein

in said register each keyword in each set of keywords
is registered in association with a predetermined
weight, and

wherein in said collator, each of said
extracted keywords is attached with a weight, and said
collator evaluates, for each form sheet type, the
degree of matching between said input form sheet and
said predetermined form sheet types by using said
weights of said extracted keywords and said
predetermined weight of each keyword in each set of
said keywords within said keyword register to thereby
decide that a form sheet type having a highest degree
of matching is the form sheet type of said input form
sheet.

21.   An apparatus according to claim 20, wherein
the weight given to each of said extracted keywords is
a weight based on a type of characters constituting the
keyword.

22.   An apparatus according to claim 21, wherein
said type of characters distinguishes whether each said
extracted keywords is typed one or handwritten one.

23.   An apparatus according to claim 20, wherein
the weight given to each of said extracted keywords is
a weight based on a location of the keyword on said
input form sheet.

24.   An apparatus according to claim 20, wherein
said predetermined weight of each keyword in each set
of keywords registered in said register is a keyword-

specific weight.

25.   An apparatus according to claim 22, wherein each of said extracted keywords is given a weight larger than 0 when the keyword is typed, and given a weight of 0 when the keyword is handwritten, such that among said extracted keywords of said input form sheet, one or more handwritten keywords are eliminated from the determination of the form sheet type.

26.   An apparatus according to claim 22, wherein the weight attached to each of said extracted keywords of said input form sheet is given a larger weight as the location of the keyword on the input form sheet approaches closer to the uppermost location.

27.   An apparatus according to claim 20, wherein the weights attached to each of said extracted keywords of said input form sheet include a weight based on the type of characters forming the keyword and a weight based on the location of the keyword on said input form sheet.

28.   An apparatus according to claim 15, further comprising a keyword forming unit which takes out arbitrary two or more keywords from keywords extracted in said keyword extracting unit, and forms one or more new keywords by combining the taken out keywords, and
        wherein said determining unit collates, for each form sheet, the extracted keywords as well as said newly formed keywords with said sets of keywords registered in said register.

29.   An apparatus according to claim 15, wherein said register includes files provided one for each form sheet type, each file registering therein a set of keywords for determining a specific form sheet.

30.   A form sheet type determining apparatus for determining to which of predetermined form sheet types an input form sheet corresponds, comprising:

a keyword register which stores a plurality of sets of keywords one set for each form sheet type;

a character recognition unit which reads image data of an input form sheet, extracts character strings from the read image data, and performs character recognition on each character string extracted;

a data base which stores reference character string pattern data;

a Keyword extraction unit which collates the character-recognized character strings with said reference character-string patterns and extracts as a keyword each of character-recognized character strings which each at least partly coincide with any of said reference character-string patterns; and

a collator which collates, for each form sheet type, said extracted keywords with said sets of keywords registered in said register, thereby to determine the type of said input form sheet.

31.   An apparatus according to claim 30, further comprising a keyword forming unit which takes out

arbitrary two or more keywords from keywords extracted in said keyword extracting unit, and forms one or more new keywords by combining the taken out keywords, and

wherein said collator collates, for each form sheet, the extracted keywords as well as said newly formed one or more keywords with said sets of keywords registered in said register. thereby to determine the type of said input form sheet.

32.   A computer program product comprising:

a computer usable medium having computer readable program code means embodied in said medium for determining whether an input form sheet is which one of predetermined form sheet types, said computer readable program code means comprising:

means for registering a plurality of sets of keywords for each of predetermined form sheet types as a set of keywords beforehand in a keyword register;

means for reading image data of input form sheet, extracting character strings from the read image data, and performing character recognition on each of the extracted character strings; and

collating means for collating, for each form sheet type, said extracted keywords with said sets of keywords registered in said keyword register, thereby to determine the type of said input form sheet.

33.   A computer program product according to claim 32, wherein in said register means, each keyword in said sets of keywords is registered in association with

a predetermined corresponding weight, and

said collating means evaluates, for each form sheet type, the degree of matching between said input form sheet and said predetermined form sheet types by using the weights given to each of said extracted keywords and said predetermined weights of the keywords in each set of said keywords within said keyword register to thereby decide that a form sheet type having a highest degree of matching is the form sheet type of said input form sheet.

34.   A computer program product according to claim 32, further comprising means for forming new keyword which takes out arbitrary two or more keywords from keywords extracted by said extracting means, and forms one or more new keywords by combining the taken out keywords, and

wherein said evaluating means includes collating means for collating, for each form sheet type, said extracted keywords and said formed one or more new keywords with said sets of keywords registered in said register.

35.   A computer program product comprising:

a computer usable medium having computer readable program code means embodied in said medium for determining whether an input form sheet is which one of predetermined form sheet types, said computer readable program code means comprising:

means for storing a plurality of sets of

keywords one set for each form sheet type;

character recognition means for reading image data of an input form sheet, extracting character strings from the read image data, and performing character recognition on each character string extracted;

keyword extraction means which collates the character-recognized character strings with said reference character-string patterns and extracts as a keyword each of character-recognized character strings which each at least partly coincide with any of said reference character-string patterns; and

collating means which collates, for each form sheet type, said extracted keywords with said sets of keywords registered in said register, thereby to determine the type of said input form sheet.

36.   A computer program product according to claim 35, wherein said computer readable program code means further comprises means for forming new keyword which takes out arbitrary two or more keywords from keywords extracted by said extracting means, and forms one or more new keywords by combining the taken out keywords, and

said collating means for collating, for each from sheet type, said extracted keywords and said formed one or more new keywords with said sets of keywords registered in said keyword register.